208 IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 28, NO. 2, FEBRUARY 2010

Distance-Enhancing Constrained Codes with Parity-Check Constraints for Data Storage Channels

Kui Cai, *Member, IEEE,* Kees A. Schouhamer Immink, *Fellow, IEEE,* Yuan Xing Lee, *Senior Member, IEEE,* Zhiliang Qin, *Member, IEEE,* and Tow Chong Chong

***Abstract*—This paper proposes efficient distance-enhancing constrained codes with parity-check (PC) constraints for data storage channels. We first propose simple and efficient finite-state encoding methods to design various distance-enhancing constrained codes, including a repeated minimum transition runlength (RMTR) code for optical recording channels, as well as a maximum transition run (MTR) code for magnetic recording channels. We further propose a general and systematic code design methodology, which can efficiently combine constrained codes with PC codes. The constrained codes can be any distance-enhancing constrained codes. The PC codes can be any linear binary PC codes. The rates of the designed codes are only a few tenths of a percent below the theoretical maximum. The proposed method enables soft information to be available to the PC decoder and soft decoding of PC codes. Examples of several newly designed distance-enhancing constrained PC codes are illustrated. Simulation results with blu-ray disc (BD) systems show that the proposed new RMTR code and RMTR constrained 4-bit PC code perform 0.2 dB and 0.85 dB better than the standard 17PP code, respectively, at error correction code (ECC) failure rate (EFR) of 10^{-12} and high recording density.**

***Index Terms*—Distance-enhancing constrained codes, error correction codes (ECCs), parity-check codes, finite-state encoding method, soft decoding, post-processor.**

I. Introduction

DEVELOPMENT of ‘efficient and powerful channel codes’ is key to ensuring good reception performance under aggressive recording conditions. There are two types of channel codes in data storage channels, namely, constrained codes [1], [2] and error correction codes (ECCs) [3]. The constrained codes, also known as modulation codes, are used to match the data to the recording channel characteristics, to improve the detector performance and to help in the operation of control loops (*e.g.* timing/gain loops) at the receiver. Runlength limited (RLL) codes [2] are examples of widely used constrained codes, which are characterized by the minimum runlength constraint d and the maximum runlength constraint k. In recent years, a maximum transition run (MTR) constraint j [4] has been further introduced to magnetic recording channels. The MTR codes prohibit input data patterns that support some of the dominant error events (*i.e.* the most likely error events that can occur) at the output of the channel detector, and therefore increase the minimum distance between those that remain. Therefore, they are known as distance-enhancing codes [5].

In optical recording systems, $d = 1$ constrained codes, examples of early distance-enhancing codes, are used for the latest blue laser disc systems (*i.e.* the blu-ray disc (BD) [6] or high-definition digital versatile disc (HD-DVD) [7]). The $d = 1$ codes are MTR codes with a $j = 1$ constraint. Furthermore, there is a second distance-enhancing constraint t [6], [7], which stipulates the maximum number of consecutive minimum distance transitions, *i.e.* runs of 2T patterns on the optical disc. The corresponding codes are referred to as repeated minimum transition runlength (RMTR) codes. For example, the standard rate 2/3 17PP code [6] used for BD is with a $t = 6$ constraint, and the eight-two-twelve-modulation (ETM) code [7] proposed for HD-DVD is with a $t = 5$ constraint. The main reason for imposing the RMTR constraint lies in the aspect that for $d = 1$ coded optical recording channels, the RMTR constraint can reduce the occurrence of consecutive 2T patterns in the channel bit-stream, which causes dominant error events at the channel detector output. In addition, it has also been found that the RMTR constraint can help to increase system tolerances, especially against tangential tilt [6], [7].

In data storage channels, the error control system typically consists of a Reed-Solomon (RS) code, which is capable of correcting combinations of random and burst errors. In recent years, soft-decodable inner ECCs have been further adopted, to make use of the soft information from the channel and effectively correct short random errors that are not covered by the outer RS-ECC. Low redundancy parity-check (PC) polynomial codes are examples of inner ECCs that are widely used in data storage channels [8], [9]. The PC code can detect the specific dominant error events of the system using only a few parity bits. For error correction, the matched-filtering type post-processor that combines syndrome and soft-decision decoding [8], [9] is widely used due to its simplicity. Alternatively, the PC codes can also be decoded iteratively by using a soft-input soft-output (SISO) channel detector (*e.g.* the BCJR detector [10] or the soft-output Viterbi detector (SOVA) [11]), as well as a sum-product algorithm (SPA) decoder [12] or its low-complexity alternatives [13]. The low-density parity-

Manuscript received 25 January 2009; revised 1 August 2009. The work of K. A. S. Immink was supported by grant *Theory and Practice of Coding and Cryptography* (NRF-CRP2-2007-03). The material in this paper was presented in part at IEEE Global Communications Conference (GLOBECOM), New Orleans, LA, November 2008.

K. Cai, Z. Qin, and T. C. Chong are with Data Storage Institute, Singapore 117608 (e-mail: Cai_Kui@dsi.a-star.edu.sg; Qin_Zhiliang@dsi.a-star.edu.sg; Chong_Tow_Chong@dsi.a-star.edu.sg).

K. A. S. Immink is with Turing Machines Inc., The Netherlands (e-mail: immink@turing-machines.com).

Y. X. Lee is with LSI Corporation, San Jose, CA 95035 USA. (e-mail: yuanxing.lee@lsi.com).

Digital Object Identifier 10.1109/JSAC.2010.100210.

check (LDPC) codes [14] are another type of inner ECCs which have the potential to approach the Shannon limit [15]. Currently, LDPC codes are being explored to replace the outer RS-ECC for data storage channels.

Many attempts have been made to efficiently combine constrained codes with PC codes [16], [17], [18], [19], [20], [9]. For example, Cideciyan *et al.* [17] designed a rate $96/104$ code which satisfies the $j = 3$ MTR constraint and a specific 2-bit PC constraint. In this scheme, between every six of the rate $16/17$ MTR codes with $j = 3$ constraint[1], two parity bits are inserted. A rate $36/36$ parity-check-preserving block encoder is further used to replace the codewords that violate the $j = 3$ constraint and ensure the PC constraint is still satisfied over the combined codeword. Although this scheme achieves a high coding efficiency, it is not sufficiently general in scope for designing codes that satisfy an arbitrary distance-enhancing constraint with any given PC constraint.

Vasic *et al.* [18] proposed a method to impose the k-constraint in the LDPC codewords through deliberate bit-flippings with the expectation that the LDPC code is able to correct both the deliberate errors and the channel errors that occur during data detection. However, for data storage channels with strong distance-enhancing constraints, such as the MTR or RMTR constraints, the number of flipped bits increases significantly compared with that in k-constrained channels. This may result in substantial performance losses and a high error floor of the LDPC codes.

Wijngaarden and Immink [19] proposed an efficient scheme that combines k-constrained codes with ECCs by interleaving constrained codes and unconstrained (parity) data. In [19], various methods are proposed to construct high-rate k-constrained codewords with unconstrained bit positions reserved for the ECC parity bits. Like the method proposed in [18], this scheme also works well only for systems with loose modulation constraints (*e.g.* the k-constraint).

For optical recording channels, although several efforts have been made to combine $d = 2$ or $d = 1$ constrained codes with PC codes [20], [9], no report has been found on the design of combined RMTR and PC codes.

In this paper, we propose efficient distance-enhancing constrained codes with PC constraints for data storage channels. We first propose novel finite-state encoding methods to design various efficient distance-enhancing constrained codes, including an RMTR code for optical recording channels, as well as an MTR code for magnetic recording channels. We further propose a general and systematic code design methodology to efficiently combine distance-enhancing constrained codes with PC codes. The rates of the designed codes are only a few tenths of a percent below the theoretical maximum. The proposed code design method enables soft decoding of PC codes.

This paper is organized as follows. In Section II, we present novel finite-state encoding methods to design an RMTR code for optical recording channels, and an MTR codes for magnetic recording channels. In Section III, we propose a general and systematic code design methodology to efficiently combine the designed distance-enhancing constrained codes

[1]At the border of two codewords, the MTR constraint is relaxed to $j = 4$.

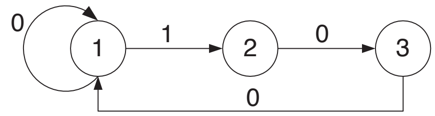


Fig. 1. FSTD for a ($d = 1$, $t = 1$) sequence.

with PC codes. The performance evaluation of the proposed RMTR code and RMTR constrained PC code is presented in Section IV. The paper is concluded in Section V.

II. Distance-Enhancing Constrained Codes

For given code constraint(s), the maximum achievable code rate was derived by Shannon and is called the Shannon Capacity [15]. The efficiency of a code, referred to as *coding efficiency*, is defined as the quotient of the code rate and Shannon Capacity. Among various code design techniques, the finite-state encoding method [1], [21] can achieve high coding efficiency, limited decoder error propagation, and satisfy strong modulation constraints, such as the RMTR or MTR constraints. Therefore, in this paper, we use this method to design the distance-enhancing constrained codes. Furthermore, unlike the *state-splitting method* [21] starting with the *labeled graph*, we propose simple and efficient finite-state encoding methods, which directly specify the encoding/decoding principles for the RMTR and MTR codes. The rates of the designed codes are only a few tenths of a percent below the capacity.

In the write path of a data storage system, a precoder of the form $1/(1 \oplus D)$ converts the binary outputs of the constrained encoder into a corresponding modulated signal, which is then stored on the storage medium. The constrained encoded bits before and after the precoder are referred to as a non-return-to-zero-inverse (NRZI) sequence, and a non-return-to-zero (NRZ) sequence, respectively. In this section, the distance-enhancing constrained codes are designed in the NRZI format.

A. RMTR Constrained Codes for Optical Recording Channels

The code design starts from computing the capacity of the RMTR codes. A finite-state transition diagram (FSTD) which provides a graphical representation of a sequence with $d = 1$ and $t = 1$ constraints is illustrated in Fig. 1. The labels on the arrows correspond to NRZI code bits, and any sequence that can be formed by following the arrows through the states and reading off the transition labels is a valid constrained sequence. An adjacency matrix for the FSTD is constructed in which the element in the i^{th} row and the j^{th} column represents the number of transitions from state i to state j. In Fig. 1, the adjacency matrix is given by $A_{(d=1,t=1)} = \begin{bmatrix} 1&1&0 \\ 0&0&1 \\ 1&0&0 \end{bmatrix}$. The capacity of the RMTR code is then computed as [15]

$$C_{(d=1,t=1)} = \log_2 \xi_{\max}[A_{(d=1,t=1)}] = 0.5515, \quad (1)$$

where $\xi_{\max}[A_{(d=1,t=1)}]$ is the largest real eigenvalue of $A_{(d=1,t=1)}$. Following a similar procedure, the capacities of various RMTR codes can be computed. They are illustrated in Table I.

From Table I, we observe that $t = 3$ is the minimum achievable RMTR constraint for $d = 1$ codes whose code

TABLE I
CAPACITY $C_{(d=1,t)}$ AS A FUNCTION OF t.

t	$C_{(d=1,t)}$
1	0.5515
2	0.6509
3	**0.6793**
4	0.6888
5	0.6922
6	0.6935
∞	0.6942

rates are comparable to that of the standard rate 2/3 codes [6], [7]. Imposing RMTR constraints stronger than $t = 3$ will introduce additional code rate loss. Therefore, in the following, we present the design of a new rate 8/12 code with the $d = 1$ and $t = 3$ constraints.

We propose a finite-state encoding method to design the new code. First, we define a codeword to be a binary string of length v that satisfies the $d = 1$ and $t = 3$ constraints. As shown in Fig. 2, the encoder has 5 states, which are characterized as follows.

- Codewords in State 1 start with ‘00’.
- Codewords in State 2 start with ‘0100’, or ‘00’.
- Codewords in State 3 start with ‘010100’, ‘0100’, or ‘00’.
- Codewords in State 4 state start with ‘100’.
- Codewords in State 5 start with ‘1010’, ‘100’, or ‘0’.

To facilitate reuse of codewords, *i.e.* mapping the same codeword to more than one user data word to achieve a high coding efficiency, each codeword may be assigned to more than one encoder state. As illustrated in Fig. 2, the rules for assigning codewords to various encoder states are as follows.

- Codewords that end with ‘00’ can be assigned to any of the five states.
- Codewords that end with ‘0010’ or ‘001010’ cannot be assigned to State 5.
- Codewords that end with ‘00101010’ or ‘001’ can be assigned to States 1 to 3.
- Codewords that end with ‘00101’ can be assigned to States 1 and 2.
- Codewords that end with ‘0010101’ can be assigned to State 1 only.

The above state-transition rules ensure that the $d = 1$ and $t = 3$ constraints are always satisfied during the concatenation of codewords from various encoder states. Furthermore, to ensure unique decodability, there is no common codeword between the five encoder states. During decoding, by observing both the current and the next codewords, the decoder can uniquely determine the user data word that was actually transmitted. Therefore, the decoder is a sliding-block decoder [1] with the least decoding window of length 2 (the minimum length possible).

Example 1: A simple example might be helpful to understand the above proposed code design method. Table II is a look-up table for a rate 3/6 RMTR code with $d = 1$ and $t = 3$ constraints, designed by using the proposed method. In the table, the first column shows the input user data words. The second to the sixth columns show the codewords mapped

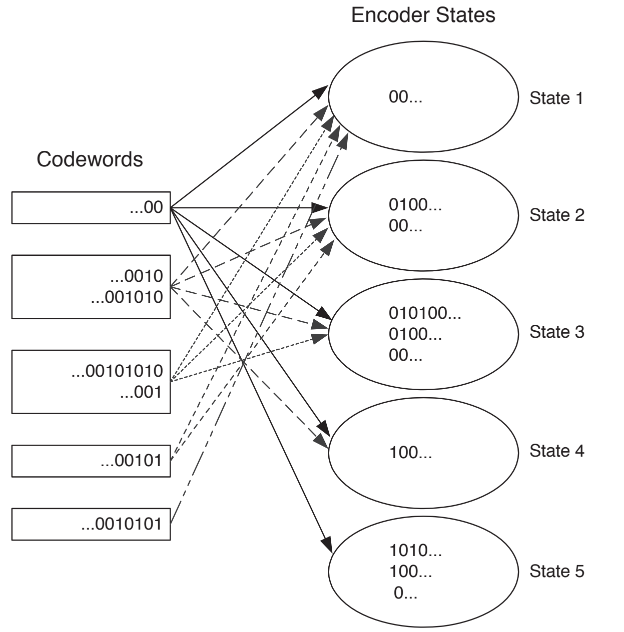


Fig. 2. Assignment of codewords to various encoder states for RMTR codes with $d = 1$ and $t = 3$ constraints.

TABLE II
CODE TABLE OF A RATE 3/6 RMTR CODE WITH $d = 1$ AND $t = 3$ CONSTRAINTS.

User data word	State 1		State 2		State 3		State 4		State 5	
	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state
000	000000	1	010000	1	010100	1	100000	1	101000	1
001	000000	2	010000	2	010100	2	100000	2	101000	2
010	000000	3	010000	3	010100	3	100000	3	101000	3
011	000000	4	010000	4	010100	4	100000	4	101000	4
100	000000	5	010000	5	010100	5	100000	5	101000	5
101	000010	1	010010	1	010001	1	100001	1	101001	1
110	000010	2	010010	2	010001	2	100001	2	101001	2
111	000010	3	010010	3	010001	3	100001	3	101001	3

to the user data words and their associated next-states, for encoder states 1 to 5, respectively. As can be seen from the table, a set of 2^u codewords, with $u = 3$ being the length of the user data words (binary), is allocated to each encoder state. Within each state, each codeword is assigned to a ‘next-state’, which stipulates the state of the encoder from which to select a codeword for the next user data word. Note that in the table, each codeword can be mapped to multiple user data words, with the corresponding next-states being different. Note also that different encoder states do not have the same codeword.

At the beginning of encoding, we assume an initial encoder state, and translate the first user data word into a codeword based on the code table. At the same time, the ‘next-state’ assigned with the current codeword indicates the state from which to select a codeword for the next data word. In this way, the encoding process is carried out sequentially. At the decoder side, we need to first look at the next codeword to obtain the encoder state that the current codeword is assigned to, and then decode the current codeword accordingly.

Following the above described finite-state encoding method, a new rate 8/12 RMTR code with 5 encoder states is designed, which satisfies the $d = 1$, $t = 3$, and $k = 16$ constraints. The rate of the code is only 1.86% below the capacity. Compared with the standard rate 2/3 codes, it imposes the minimum

achievable RMTR constraint on the channel bit-stream with the least decoding window length, without introducing additional code rate loss.

B. MTR Constrained Codes for Magnetic Recording Channels

In this section, we propose efficient MTR constrained codes designed for magnetic recording channels. We use the $j = 3$ MTR codes as an example, although the code design method can be generalized to other MTR codes, such as the $j = 4$ MTR codes as well.

The technique of look-ahead coding in conjunction with violation detection and substitution is widely use for constructing MTR codes [5]. In this section, however, we propose an efficient finite-state encoder which transforms u-bit user data words into v-bit codewords that satisfy the $j = 3$ constraint. In the proposed method, to better allocate the valid codewords to different encoder states to achieve a high coding efficiency, we classify the various encoder states into state sets of different types. As shown in Fig. 3, the encoder has s states, which are classified into four state sets of a first, second, third, and fourth type. The encoder has s_1 states of the first type, s_2 states of the second type, s_3 states of the third type, and $s_4 = s - s_1 - s_2 - s_3$ states of the fourth type. The four types of encoder states are characterized as follows.

- Codewords in states of the first type start with '0'.
- Codewords in states of the second type start with either '0', or '10'.
- Codewords in states of the third type start with '0', '10', or '110'.
- Codewords in states of the fourth type start with '0', '10', '110', or '1110'.

The state-transition rules are as follows.

- Codewords that end with a '0' can be assigned to any of the s encoder states.
- Codewords that end with a '01' cannot be assigned to the s_4 states of the fourth type.
- Codewords that end with a '011' can be assigned to the s_1 states of the first type, or the s_2 states of the second type.
- Codewords that end with a '0111' can be assigned to the s_1 states of the first type only.

The above state-transition rules avoid violation of the $j = 3$ constraint while concatenating codewords from various encoder states.

Furthermore, as in the case of designing the RMTR codes, due to the reuse of codewords in encoding, different encoder states (of any type) cannot have the same codeword. This attribute implies that any codeword can be unambiguously related to the state from which it emerged. That is, similar to the case with RMTR codes, the proposed decoder is a sliding-block decoder with one codeword look-ahead.

With the above described finite-state encoding method, we obtain the following conditions:

$$
\begin{aligned}
& s|E_{0\cdots 0}| + (s_1+s_2+s_3)|E_{0\cdots 01}| \\
& +(s_1+s_2)|E_{0\cdots 011}| + s_1|E_{0\cdots 0111}| \geq s_1 2^u, \qquad (2) \\
& s\,(|E_{0\cdots 0}| + |E_{10\cdots 0}|) + (s_1+s_2+s_3)(|E_{0\cdots 01}| \\
& +|E_{10\cdots 01}|) + (s_1+s_2)(|E_{0\cdots 011}| + |E_{10\cdots 011}|) \\
& +s_1\,(|E_{0\cdots 0111}| + |E_{10\cdots 0111}|) \geq (s_1+s_2)2^u, \qquad (3) \\
& s\,(|E_{0\cdots 0}| + |E_{10\cdots 0}| + |E_{110\cdots 0}|) \\
& +(s_1+s_2+s_3)(|E_{0\cdots 01}| + |E_{10\cdots 01}| + |E_{110\cdots 01}|) \\
& +(s_1+s_2)(|E_{0\cdots 011}| + |E_{10\cdots 011}| + |E_{110\cdots 011}|) \\
& +s_1\,(|E_{0\cdots 0111}| + |E_{10\cdots 0111}| + |E_{110\cdots 0111}|) \\
& \qquad \geq (s_1+s_2+s_3)2^u, \qquad (4) \\
& s\,(|E_{0\cdots 0}| + |E_{10\cdots 0}| + |E_{110\cdots 0}| + |E_{1110\cdots 0}|) \\
& +(s_1+s_2+s_3)(|E_{0\cdots 01}| + |E_{10\cdots 01}| + |E_{110\cdots 01}| \\
& +|E_{1110\cdots 01}|) + (s_1+s_2)(|E_{0\cdots 011}| + |E_{10\cdots 011}| \\
& +|E_{110\cdots 011}| + |E_{1110\cdots 011}|) + s_1(|E_{0\cdots 0111}| \\
& +|E_{10\cdots 0111}| + |E_{110\cdots 0111}| + |E_{1110\cdots 0111}|) \geq s2^u, \qquad (5)
\end{aligned}
$$

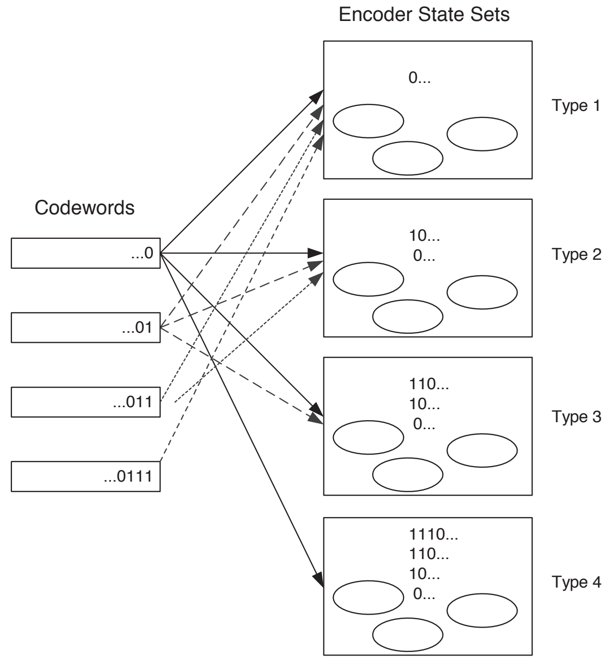


Fig. 3. Assignment of codewords to encoder state sets of different types for MTR codes with a $j = 3$ constraint.

where $E_{a\cdots b}$ denotes the set of codewords starting with a binary string of 'a' and ending with a binary string of 'b', and $|E_{a\cdots b}|$ denotes the cardinality of the codeword set $E_{a\cdots b}$. Note that the above inequalities are equivalent to the *approximate eigenvector inequality* [21], and they are necessary conditions for the code construction. By using computer search, we can determine suitable integers s, s_1, s_2, and s_3 which satisfy conditions of (2) to (5) to maximize the code rate u/v, and design the code accordingly.

By using the above described code design method, a new rate 17/18 MTR code with 8 encoder states ($s = 8$, $s_1 = 4$, $s_2 = 2$, $s_3 = 1$, and $s_4 = 1$) is designed, which satisfies the $j = 3$ and $k = 13$ constraints. Note that the capacity of $j = 3$ MTR codes is 0.9468 [5]. Note also that the code rate of the state of the art $j = 3$ MTR codes is 16/17 [5], [22]. Therefore, with a coding efficiency of 99.75%, the proposed new MTR code achieves a higher efficiency than the state of the art $j = 3$ MTR codes.

III. Distance-Enhancing Constrained Codes With PC Constraints

To further impose the PC constraints on the designed RMTR and MTR codes to achieve a higher coding gain, we propose a general and systematic method to efficiently combine distance-enhancing constrained codes with PC codes. The obtained codes are referred to as the RMTR or MTR constrained PC codes.

A. Capacity of Constrained PC Codes

When the PC constraints are imposed on the channel bit-stream, the modulation constraints should be satisfied simultaneously. This will result in additional code rate loss. The minimum overhead is one user bit per parity bit. Equivalently, $\frac{1}{R_{Cap}}$ channel bits are needed per parity bit, where R_{Cap} is the capacity of the constrained code. Let there be p parity bits per codeword of length n. Then, the capacity of constrained PC codes is given by

$$C_{PC} \quad = \quad R_{Cap} - \frac{p}{n}. \tag{6}$$

B. General Principle of the New Code Design Methodology

For a sequence of distance-enhancing constrained codewords, we propose to further design a specific parity-related constrained (PRC) code to realize certain PC constraint over the combined codeword, which is a concatenation of the sequence of distance-enhancing constrained codewords and the PRC codeword. The PC constraint corresponds to a predetermined generator matrix (or generator polynomial) of a linear binary PC code [3]. For ease in imposing the modulation constraints, the generator matrix needs to be designed to generate a systematic code. The proposed code design method is based on the following observation.

- *Consider an $[l, n]$ systematic linear binary PC code C with $p = l - n$ parity bits. Let $\mathbf{v_1}$ and $\mathbf{v_2}$, respectively, denote row vectors with n_1 bits and $n_2 = n - n_1$ bits consisting of a sequence of distance-enhancing constrained codewords and a PRC codeword, with $0 < n_1 < n$. If the parity bits of* $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{0,\cdots,0}_{n_2}]$ *and* $[\underbrace{0,\cdots,0}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2}]$ *are equal, then the combined constrained codeword* $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2}]$*, with p bits of zeros appended, generates a codeword of C.*

Let $\mathbf{G} = [\mathbf{I}\ \mathbf{P}]$ be a generator matrix of C, where $\mathbf{I}$ is an $n \times n$ identity matrix, and $\mathbf{P}$ is an $n \times p$ matrix. The parity bits of $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{0,\cdots,0}_{n_2}]$ and $[\underbrace{0,\cdots,0}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2}]$ are computed as

$$\mathbf{p_1} = [\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{0,\cdots,0}_{n_2}]\mathbf{P} \text{ and } \mathbf{p_2} = [\underbrace{0,\cdots,0}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2}]\mathbf{P}. \tag{7}$$

If $\mathbf{p_1} = \mathbf{p_2}$, we get $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2}]\mathbf{P} = [\underbrace{0,\cdots,0}_{p}]$. Thus, $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2} \mid \underbrace{0,\cdots,0}_{p}]$ is a codeword of C.

Therefore, the structure of the combined constrained PC code includes two component codes: the distance-enhancing

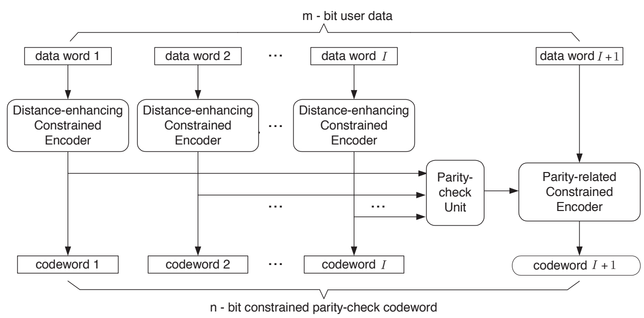


Fig. 4. Block diagram for encoding distance-enhancing constrained codes with PC constraints.

constrained code and the PRC code. During encoding, as shown in Fig. 4, an m-bit segment of user data is partitioned into $I + 1$ data words. The I leading data words are first encoded into distance-enhancing constrained codewords. The parity bits $\mathbf{p_1}$ associated with this sequence of codewords are then computed. After that, a specific PRC codeword, whose associated parity bits $\mathbf{p_2}$ are the same with $\mathbf{p_1}$, is selected from a candidate codeword set and concatenated directly with the distance-enhancing constrained codewords, thus forming the combined constrained PC codeword $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2}]$. It is essential that the distance-enhancing constrained code and the PRC code should be constructed based on the same finite-state encoding method, such as those proposed in Section II. This enables the two component codes to be connected in any order without violating the modulation constraints. The details for designing the PRC code will be presented in the next section.

Remarks:

- In this work, for simplicity, $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2}]$ is referred to as the constrained PC codeword. However, $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2} \mid \underbrace{0,\cdots,0}_{p}]$ is actually the complete codeword of the PC code C.
- According to (7), during encoding, to compute the parity bits $\mathbf{p_1}$ and $\mathbf{p_2}$ associated with $\mathbf{v_1}$ and $\mathbf{v_2}$, n_2 trailing zeros and n_1 leading zeros need to be appended to $\mathbf{v_1}$ and $\mathbf{v_2}$, respectively.

The combined constrained PC codeword is then transmitted over the channel without appending its parity bits $[\underbrace{0,\cdots,0}_{p}]$, since the latter is fixed and known by the receiver. However, at the receiver side, the decoder of C needs to have the complete PC codeword to do decoding. According to the description above, $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2} \mid \underbrace{0,\cdots,0}_{p}]$ is the complete codeword of C. Therefore, the parity bits $[\underbrace{0,\cdots,0}_{p}]$ need to be appended back to the received constrained codeword $[\underbrace{\hat{\mathbf{v}}_\mathbf{1}}_{n_1} \mid \underbrace{\hat{\mathbf{v}}_\mathbf{2}}_{n_2}]$ to form the reconstructed codeword $[\underbrace{\hat{\mathbf{v}}_\mathbf{1}}_{n_1} \mid \underbrace{\hat{\mathbf{v}}_\mathbf{2}}_{n_2} \mid \underbrace{0,\cdots,0}_{p}]$ of C for decoding. For the SPA decoder or its low-complexity

alternatives, the *a priori* information of $[\underbrace{\hat{\mathbf{v}}_1}_{n_1} \mid \underbrace{\hat{\mathbf{v}}_2}_{n_2}]$ can be obtained from a SISO channel detector. The *a priori* probability of each parity bit being a '0' is set to 1, since the parity bits are forced to be $[\underbrace{0, \cdots, 0}_{p}]$ by the proposed method during encoding. In this way, we reduce the redundancy that is added into the channel bit-stream and enable soft decoding of the PC code C. Note that, in principle, we could always choose $\mathbf{p_1} + \mathbf{p_2} = \mathbf{a}$, where $\mathbf{a}$ is an arbitrary p-bit row vector, and generate a codeword of C in terms of $[\underbrace{\mathbf{v_1}}_{n_1} \mid \underbrace{\mathbf{v_2}}_{n_2} \mid \underbrace{\mathbf{a}}_{p}]$.

The rate of the new constrained PC code is given by

$$R = \frac{m}{n} = R_1 - \frac{n_2}{n}(R_1 - R_2), \tag{8}$$

where m and n are the lengths of the segment of user data and the combined constrained PC codeword, respectively, and R_1 and R_2 are the rates of the distance-enhancing constrained code and the PRC code, respectively.

C. Design of the PRC Codes

To design the PRC code, as described in the previous section, we use the same finite-state encoding method as for the distance-enhancing constrained code. However, for the distance-enhancing constrained code, to achieve a high coding efficiency, there is only one codeword mapped to each user data word, in each of the encoder states. For the PRC code, on the other hand, there is a set of 2^p candidate codewords potentially mapped to each user data word, in each of the encoder states, where p is the number of parity bits. Therefore, compared with the distance-enhancing constrained code with the same input user data word length, the codeword length of the PRC code needs to be longer to satisfy the additional PC constraint, and hence its code rate is smaller. Therefore, to achieve a high coding efficiency, as shown in Fig. 4, each combined constrained PC code consists of multiple distance-enhancing constrained codewords. The PRC codeword is only used as the last codeword in the combined codeword.

We propose the following criteria that guides the design of the PRC code.

- *To design a PRC code with m_2 user data bits and p parity bits, there should be at least 2^{m_2+p} codewords allocated to each of the encoder states (of any type). Furthermore, for each set of codewords with the same parity bits, there should be at least 2^{m_2} codewords allocated to each encoder state (of any type).*

The rate of the PRC code is then given by

$$R_2 = m_2/n_2. \tag{9}$$

The main steps for the design of the PRC code are as follows.

1) For a PRC code with m_2 user data bits and p parity bits, use the criteria described above to determine n_2 and a suitable number of encoder states.
2) Enumerate all the valid constrained codewords of length n_2. Based on the given generator matrix, compute the parity bits $\mathbf{p_2}$ of each codeword and distribute them into a group of codeword sets. A total of 2^p codeword sets are obtained.
3) For each set of codewords with the same parity bits, allocate the codewords to various encoder states by following the encoding method of the distance-enhancing constrained code. This results in a set of 2^p sub-tables.
4) Concatenate the 2^p sub-tables together, and form a code table for encoding/decoding of the PRC code. Compared with the code table of the distance-enhancing constrained code, the PRC code table is enlarged by a factor of 2^p. In each encoder state, there is a set of 2^p codewords potentially mapped to one user data word.

TABLE III
CODE TABLE OF A RATE 3/8 PRC CODE WITH $d = 1$ AND $t = 3$ CONSTRAINTS.

Parity even										
User data word	State 1		State 2		State 3		State 4		State 5	
	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state
000	00000000	1	01000100	1	01010000	1	10000100	1	10100000	1
001	00000000	2	01000100	2	01010000	2	10000100	2	10100000	2
010	00000000	3	01000100	3	01010000	3	10000100	3	10100000	3
011	00000000	4	01000100	4	01010000	4	10000100	4	10100000	4
100	00000000	5	01000100	5	01010000	5	10000100	5	10100000	5
101	00001001	1	01000010	1	01000001	1	10000001	1	10000010	1
110	00001001	2	01000010	2	01000001	2	10000001	2	10000010	2
111	00001001	3	01000010	3	01000001	3	10000001	3	10000010	3

Parity odd										
User data word	State 1		State 2		State 3		State 4		State 5	
	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state	Codeword	Next-state
000	00000100	1	01000000	1	01010010	1	10000000	1	10100100	1
001	00000100	2	01000000	2	01010010	2	10000000	2	10100100	2
010	00000100	3	01000000	3	01010010	3	10000000	3	10100100	3
011	00000100	4	01000000	4	01010010	4	10000000	4	10100100	4
100	00000100	5	01000000	5	01010001	1	10000000	5	10100100	5
101	00000001	1	00000010	1	01001001	1	10001001	1	10100010	1
110	00000001	2	00000010	2	01001001	2	10001001	2	10100010	2
111	00000001	3	00000010	3	01001001	3	10001001	3	10100010	3

Example 2: We use the following simple example to illustrate the process to design the PRC code. Assume the design of a PRC component code with $m_2 = 3$, for an RMTR constrained single-bit even PC code. The distance-enhancing constrained code is assumed to be the rate 8/12 RMTR code with 5 encoder states as proposed in Section II-A. First, by using the criteria proposed for designing the PRC code, we determine a minimum codeword length of $n_2 = 8$. Next, generate all the valid $d = 1$ and $t = 3$ codewords of length $n_2 = 8$ and allocate them to various encoder states according to the principles described above. An example of the code table for the resulting rate $3/8$ PRC code is illustrated by Table III. As can be seen, the code table includes two sub-tables, which contain codewords with even and odd parities, respectively. In each sub-table, the characteristics of encoder states as well as the state-transitions rule are the same as those of the rate 8/12 code. Therefore, the rate $3/8$ PRC code can be connected directly with the rate 8/12 code and the $d = 1$ and $t = 3$ constraints are preserved at the boundary of the two codes. Furthermore, in each of the encoder states, there is a set of two codewords mapped to one user data word. This ensures that during encoding, a suitable PRC codeword from the codeword set can always be chosen, to be concatenated with the sequence of RMTR codewords and to realize an even PC constraint over the combined codeword.

Based on the same code design method, the operation of the PRC decoder is generally the same as that of the distance-enhancing constrained decoder, but with the code tables being

different. Both decoders are sliding-block decoders with a decoding window of length 2.

D. Code Design in NRZ Format

In the previous section, we presented the code design method in NRZI format. For PC codes and post-processing based detection approaches, it is preferable to encode the data in NRZ format due to the reason that in the NRZI case, error detection and post-processing have to be done at the output of the 'NRZ to NRZI inverse precoder'. The process of inverse precoding will cause error propagation and thus increase the length of error events.

Several approaches have been proposed to detect and correct errors in NRZ format [8], [17]. However, they either considerably weaken the modulation constraint of the encoded channel bit-stream [8], or have some specific requirement on the PC code [17]. We now present a new approach to design the constrained PC code in NRZ format. In our approach, the code table of the distance-enhancing constrained code remains the same as that in NRZI format. However, the code table for the PRC code is designed in a different way. The details are as follows. First, determine the codeword length n_2 and the suitable number of encoder states for the PRC code follow criteria similar to those in the NRZI case. The only difference is that the parity bits of each codeword are computed in the NRZ format, rather than in the NRZI format, based on an assumed initial NRZ bit. For example, in Table III, the codeword '10000001' has an even parity in the NRZI format. However, with an initial NRZ bit of '0', the codeword in NRZ format (*i.e.* '11111110') has an odd parity instead. Second, enumerate all the valid codewords of length n_2 in NRZI format. Compute the parity bits of the codewords in NRZ format with an assumed initial NRZ bit as described above. Third, distribute each set of NRZI codewords with the same NRZ parity bits into different encoder states, and form a set of 2^p sub-tables. Fourth, concatenate the 2^p sub-tables together to form the code table for the PRC code.

To do encoding, the distance-enhancing constrained codewords are first constructed and connected as in the NRZI case. The resulting codewords are then converted into NRZ format by a precoder, and the associated parity bits are computed. Based on these parity bits as well as the last bit of the NRZ sequence, the PRC codeword in NRZI format that has the same NRZ parity bits is selected from the codeword set. The PRC codeword needs to be converted into NRZ format before concatenating with the NRZ format distance-enhancing constrained codewords. During decoding, the detected NRZ data sequence is first converted into NRZI format through an inverse precoder, and the resulting NRZI sequence is then decoded based on the code tables of the distance-enhancing constrained code and the PRC code.

E. Examples of Newly Designed Codes

In this section, we present several efficient distance-enhancing constrained codes with PC constraints for optical and magnetic recording channels, which are designed by using the above proposed code design method. The codes are designed in both NRZI and NRZ formats.

1) RMTR Constrained PC Codes: Using the new rate 8/12 code designed in Section II-A as one component code, various RMTR constrained PC codes that satisfy the $d = 1$ and $t = 3$ constraints can be designed. We exemplify our code design with a new constrained 2-bit PC code and a new constrained 4-bit PC code. They are defined by generator polynomials $g(x) = 1+x+x^2$ and $g(x) = 1+x+x^4$, respectively. A rate 8/15 ($d = 1$, $t = 3$, and $k = 16$) code and a rate 8/18 code ($d = 1$, $t = 3$, and $k = 16$) are designed as the PRC codes, respectively. Similar to the rate 8/12 code, the encoders of both codes have 5 states. The rates of the combined constrained PC codes are 136/207 and 272/414, respectively. Both codes achieve a coding efficiency of only 1.84% below the capacity.

Note that for the proposed constrained PC codes, the lengthes of the input user data words for both the RMTR and PRC component codes are designed to be 8 bits. Therefore, if these codes are used in conjunction with an outer RS-ECC with the same symbol size (of 8-bit), error propagation due to the mismatch between the input user data words length of the constrained code and the symbol size of RS-ECC is avoided. The MTR constrained PC codes proposed in the next section are designed in a similar fashion.

2) MTR Constrained PC Codes: Based on the new rate 17/18 code proposed in Section II-B, various $j = 3$ constrained PC codes can be designed. The following 2-bit and 4-bit MTR constrained PC codes are two examples. The PRC codes are a new rate 17/21 ($j = 3$, $k = 13$) code and a new rate 17/23 ($j = 3$, $k = 13$) code corresponding to generator polynomials $g(x) = 1 + x + x^2$ and $g(x) = 1 + x + x^4$, respectively. The encoders of both codes have 8 states ($s = 8$, $s_1 = 4$, $s_2 = 2$, $s_3 = 1$, and $s_4 = 1$). The rates of the combined 2-bit and 4-bit constrained PC codes are 136/147 and 255/275, respectively. The corresponding coding efficiencies are 99.21% and 99.44%, respectively. In particular, the coding efficiency of the rate 136/147 2-bit constrained PC code is higher than that of the rate 96/104 2-bit PC code proposed in [17].

Finally, we remark that the proposed code design method can be used to efficiently combine distance-enhancing constrained codes with inner ECCs, such as the PC polynomial codes illustrated above, or LDPC codes which are based on multiple single-bit PC codes (*e.g.* those proposed in [23]). From practical implementation viewpoint, the proposed method cannot be directly applied to outer RS-ECCs which have hundreds to thousands parity bits, since in this case a prohibitive large number of look-up tables (*i.e.* 2^p sub-tables) are needed for the PRC code. Effective coding techniques, such as the enumeration method might be further used to replace table look-up. However, such explorations are beyond the scope of the current paper.

IV. Performance Evaluation

The performance of the MTR codes have been widely studied for magnetic recording channels [4], [22], [17], [24], [25]. In particular, the performance improvement of the rate 16/17 $j = 3$ MTR code has been reported in [22], [17], [24]. In [17], the performance gain achieved by the rate 96/104 $j = 3$ MTR constrained PC code with 2 parity-bit has also

been reported. Moreover, in [24], it has been found that the $j = 3$ MTR codes can help to delay the error rate floor of LDPC codes, through eliminating the long error events of the channel. Since the code rates of the $j = 3$ MTR code proposed in Section II-B and $j = 3$ MTR constrained 2-bit PC code proposed in Section III-E2 are slightly higher than those of the rate 16/17 MTR code and the rate 96/104 code, respectively, they show a high potential for the application in magnetic recording channels. For optical recording channels, however, no much report has been found on the performance of RMTR codes and RMTR constrained PC codes. Therefore, in this section, we focus on evaluating the performance of the proposed new codes for optical recording channels.

In the following, performance of the proposed new RMTR code and RMTR constrained 4-bit PC code is evaluated for high-density BD systems. In particular, the performance of the new codes is compared with the standard 17PP code using the ECC failure rate (EFR) at the output of the ECC decoder as the performance criterion. We focus on a RS [248, 216] code with 8-bit/symbol (byte), since it serves as the key constituent of the ECC configuration for BD [6].

In this work, we assume an idealized BD system and use a generalized Braat-Hopkins model [26] to describe the channel. In the model, the quantity $\Omega_u = f_c T_u$, which is the optical cut-off frequency f_c normalized by user bit rate $1/T_u$, is a measure of the recording density. For BD systems with high density, we get $\Omega_u = 0.39$ [9]. At the channel side, $\Omega_c = f_c T_c$ is the channel bit rate normalized cut-off frequency with $T_c = RT_u$. Here, R is the net code rate arising from all the channel codes, *i.e.* constrained codes and ECCs. During the performance comparison among different codes, we always keep the user density (in terms of Ω_u) the same and adjust the channel density (in terms of Ω_c) to account for the influence of different code rates. The variance σ^2 of additive white Gaussian channel noise, is determined by the user signal-to-noise-ratio (SNR) defined as $\mathrm{SNR_u(dB)} = 10\log_{10}\left(\frac{\sum h_{ku}^2}{\sigma_u^2}\right)$ and $\sigma^2 = \frac{1}{R}\sigma_u^2$, where σ_u^2 is the noise power in the user bandwidth, and h_{ku} is the channel symbol response for $R = 1$ and $\Omega_u = 0.33$. These definitions of channel response and SNR help to fairly reflect the impact of code rate in the performance evaluation. They are believed to reflect reasonably the actual physical storage system.

In the simulations, a Viterbi detector (VD) that is matched to a 7-tap optimized partial response (PR) target is used as the channel detector [9]. Fig. 5 shows the histograms of dominant error events at the VD output, for the case with 17PP code. The probabilities of the first five dominant error events, normalized by the error event rate, are particularly indicated for various SNR values. Observe that the first five dominant error events turn out to be $\pm\{2\}$, $\pm\{2,0,-2\}$, $\pm\{2,0,-2,0,2\}$, $\pm\{2,0,-2,0,2,0,-2\}$, and $\pm\{2,0,-2,0,2,0,-2,0,2\}$. Moreover, the remaining less dominant error events (denoted by 'other events') are $\pm\{2,0,-2,0,2,0,-2,0,2,0,-2\}$, $\pm\{2,0,-2,0,2,0,-2,0,2,0,-2,0,2\}$, $\pm\{2,0,-2,0,2,0,-2,0,2,0,-2,0,2,0,-2\}$, *etc*. Obviously, except for the single-bit error event $\pm\{2\}$, all the above error events are caused by consecutive 2T patterns. The above analysis coincides with the observations in [6], [7].

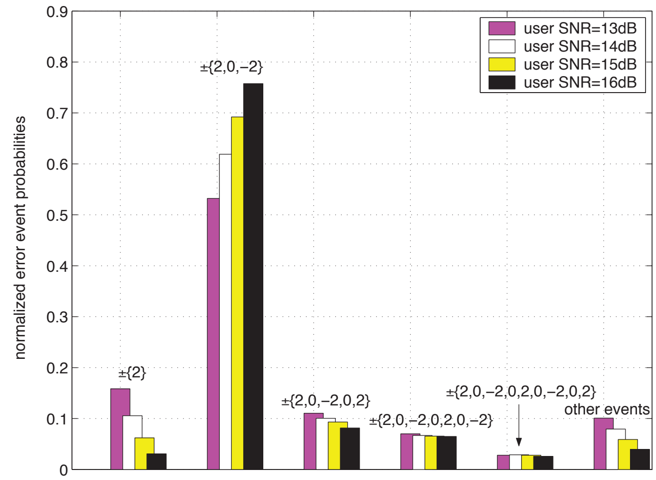


Fig. 5. Histogram of dominant error events at VD output, with BD systems, 17PP code, and $\Omega_u = 0.39$.

In the simulations with RMTR codes, all the states and state transitions which correspond to invalid data patterns are removed from the VD trellis. For the case with PC, a matched-filtering type post-processor is used at the output of the VD, which can correct both single and double error events that occur within each detected codeword [9]. It can be verified that all the five dominant error events illustrated above can be detected by the proposed RMTR constrained 4-bit PC code. To correct these error events, five matched-filters need to be used in the post-processor for the the case with $d = 1$ constrained PC code [9]. However, with the proposed RMTR constrained PC code, only three filters that match to the events $\pm\{2\}$, $\pm\{2,0,-2\}$, $\pm\{2,0,-2,0,2\}$ are needed, since the remaining dominant error events are prohibited by the $t = 3$ constraint. This helps to reduce the complexity of the post-processor.

In this work, we use semi-analytical approaches proposed in [27] to estimate the EFR of the RS code with various constrained codes. In particular, we apply the multinomial and block multinomial methods to estimate the failure rates of RS-ECC, for the cases of without and with PC code, respectively. Interleaving is a useful tool to enhance the capability of ECC for correcting burst errors [3], which have been found to be dominant in high-density optical recording systems. Therefore, compared with ECCs of magnetic recording channels, the ECCs of optical recording channels have a much larger interleaving degree [6]. Furthermore, it has been found that for the PC-coded optical recording system, interleaving can effectively break long byte errors which are mainly caused by the mis-corrections of the post-processor into short byte errors over different interleaves. This enhances the error correction capability of the RS-ECC, and results in a larger gain for the constrained PC codes [27]. In this work, we take an interleaving degree of $\alpha = 34$, since it maximizes the EFR of the RMTR constrained 4-bit PC code.

In Fig. 6, we present the EFR comparison between the newly designed codes and the 17PP code. Observe that at $\mathrm{EFR} = 10^{-12}$, the rate 8/12 RMTR code performs around 0.2 dB better than the 17PP code, since $\pm\{2,0,-2,0,2,0,-2\}$, $\pm\{2,0,-2,0,2,0,-2,0,2\}$, and so on error events are prohibited by the $t = 3$ constraint. The performance gain

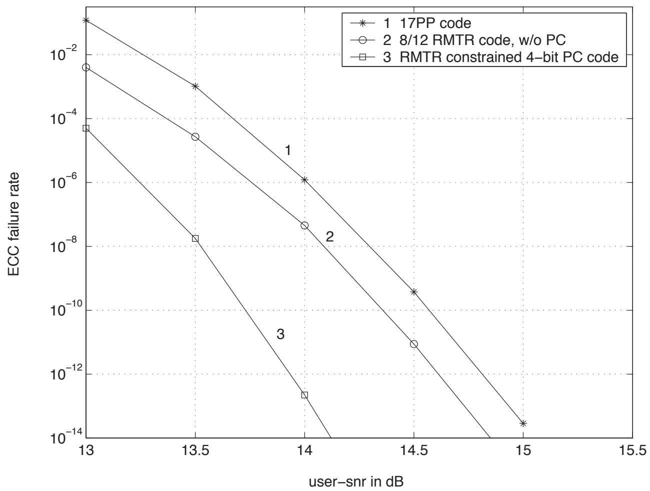


Fig. 6. EFR comparison between the new codes and the 17PP code, with BD systems and $\Omega_u = 0.39$.

for the case with PC is more significant. In particular, the RMTR constrained 4-bit PC code gains 0.85 dB over the 17PP code. This is due to the reason that the $t = 3$ constraint can completely eliminate the dominant error events $\pm\{2, 0, -2, 0, 2, 0, -2\}$ and $\pm\{2, 0, -2, 0, 2, 0, -2, 0, 2\}$, and thereby avoid the mis-corrections of these error events in the post-processor. Furthermore, the RMTR constraint can also eliminate many other long error events, such as $\pm\{2, 0, -2, 0, 2, 0, -2, 0, 2, 0, -2\}$, $\pm\{2, 0, -2, 0, 2, 0, -2, 0, 2, 0, -2, 0, 2\}$, $\pm\{2, 0, -2, 0, 2, 0, -2, 0, 2, 0, -2, 0, 2, 0, -2\}$, *etc*. These long error events require more parity bits for error detection, and introduce many more errors due to the mis-corrections of the post-processor.

V. Conclusions

In this paper, we have proposed efficient constrained codes that satisfy the distance-enhancing constraints and PC constraints for data storage channels. In particular, we first proposed simple and efficient finite-state encoding methods to design various distance-enhancing codes, including an RMTR code for optical recording channels, and an MTR codes for magnetic recording channels. Compared with the codes used in standard blue laser disc systems, the new RMTR code imposes the minimum achievable RMTR constraint on the channel bitstream, without introducing additional code rate loss. The new MTR code achieves a coding efficiency of 99.75%, which is higher than that of the state of the art MTR codes with the same j-constraint. The decoders of both codes are sliding-block decoders with the least decoding window length.

To further impose the PC constraints on the designed RMTR and MTR codes to achieve a higher coding gain, we proposed a general and systematic code design methodology to efficiently combine distance-enhancing constrained codes with PC codes. In particular, we proposed to further design a PRC code to realize the PC constraint over the combined codeword, which is a concatenation of the sequence of distance-enhancing constrained codewords and the PRC codeword. Approaches have been proposed to design the constrained PC codes either in NRZI or NRZ format. The rates of the designed codes are only a few tenths of a percent below the capacity. The proposed method enables soft information to be available to the PC decoder and soft decoding of PC codes. Examples of several newly designed distance-enhancing constrained PC codes have been illustrated. Simulation results with BD systems show that the proposed new RMTR code and RMTR constrained 4-bit PC code achieve performance gain of 0.2 dB and 0.85 dB over the standard 17PP code, respectively, at EFR $= 10^{-12}$ and high recording density.

References

[1] B.H. Marcus, P.H. Siegel, and J.K. Wolf, "Finite-state modulation codes for data storage," *IEEE Trans. Commun.*, vol. 10, no. 1, pp. 5-37, Jan. 1992.

[2] K.A.S. Immink, *Codes for Mass Data Storage Systems*. Shannon Foundation Publishers, 1999, Chaps. 4, 5, 13.

[3] S. Lin and D.J. Costello, *Error Control Coding Fundamentals and Applications*. Prentice-Hall Inc., 1983.

[4] J. Moon and B. Brickner, "Maximum transition run codes for data storage systems," *IEEE Trans. Magn.*, vol. 32, pp. 3992-3994, Sep. 1996.

[5] R.D. Cideciyan, E. Eleftheriou, B.H. Marcus, and D.S. Modha, "Maximum transition run codes for generalized partial response channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 4, pp. 619-634, Apr. 2001.

[6] T. Narahara, S. Kobayashi, Y. Shimpuku, G. van den Enden, J. Kahlman, M. van Dijk, and R. van Woudenberg, "Optical disc system for digital video recording," *Jpn. J. Appl. Phys.*, pt. 1, vol. 39, no. 2B, pp. 912-919, 2000.

[7] K. Kayanuma, C. Nota, and T. Iwanaga, "Eight to twelve modulation code for high density optical disk," in *Tech. Dig. Intl. Symp. Optical Memory* (ISOM), Nara, Japan, Nov. 2003, pp. 160-161.

[8] R.D. Cideciyan, J.D. Coker, E. Eleftheriou, and R.L. Galbraith, "Noise-predictive maximum likelihood detection combined with parity-based post-processing," *IEEE Trans. Magn.*, vol. 37, no. 2, pp. 714-720, Mar. 2001.

[9] K. Cai, K.A.S. Immink, J.W.M. Bergmans, and L.P. Shi, "Novel constrained parity-check code and post-processor for advanced blue laser disk," in *Tech. Dig. Int. Symp. Optical Media and Optcal Data Storage* (ISOM/ODS), Hawaii, USA, Jul. 2005, pp. TuB5.

[10] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, pp. 284-287, Mar. 1974.

[11] J. Hagenauer and P. Hoeher, "A Viterbi algorithm with soft decision outputs and its applications," in *Proc. IEEE Intl. Conf. Global Telecommun.* (GLOBECOM), Dallas, USA, Nov. 1989, pp. 1680-1688.

[12] X-Y. Hu, E. Eleftherious, D-M. Arnold, and A. Dholakia, "Efficient implementation of the sum-product algorithm for decoding LDPC codes," in *Proc. IEEE Intl. Conf. Global Telecommun.* (GLOBECOM), San Antonio, USA, Nov. 2001, pp. 1036-1036E.

[13] J. Chen, A. Dholakia, E. Eleftheriou, M.P.C. Fossorier, and X. Hu, "Reduced-complexity decoding of LDPC codes," *IEEE Trans. Commun.*, vol. 53, pp. 1288-1299, Aug. 2005.

[14] R.G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.

[15] C.E. Shannon,"A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379-423, 1948.

[16] J. Campello de Souza, B.H. Marcus, R. New, and B.A. Wilson, "Constrained systems with unconstrained positions," *IEEE Trans. Inf. Theory*, vol. 48, pp. 866-879, Apr. 2002.

[17] R.D. Cideciyan and E. Eleftheriou, "Codes satisfying maximum transition run and parity-check constraints," in *Proc. IEEE Intl. Conf. Commun.* (ICC), Paris, France, Jun. 2004, pp. 635-639.

[18] B. Vasic and K. Pedagani, "Run-length-limited low-density parity-check codes based on deliberate error insertion," *IEEE Trans. Magn.*, vol. 40, no. 3, pp. 1738-1743, May 2004.

[19] A.J. van Wijngaarden and K.A.S. Immink, "Maximum runlength limited codes with error control capabilities," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 4, pp. 602-611, Apr. 2001.

[20] W.M.J. Coene, H. Pozidis, and J.W.M. Bergmans, "Run-length limited parity-check coding for transition-shift errors in optical recording," in *Proc. IEEE Intl. Conf. Global Telecommun.* (GLOBECOM), San Antonio, USA, Nov. 2001, pp. 2982-2986.

[21] R.L. Adler, D. Coppersmith, and M. Hassner, "Algorithms for sliding block codes: An application of symbolic dynamics to information theory," *IEEE Trans. Inf. Theory*, vol. 29, pp. 5-22, Jan. 1983.

[22] T. Nishiya, K. Tsukano, T. Hirai, S. Mita, and T. Nara, "Rate 16/17 maximum transition run (3;11) code on an EEPRML channel with an error-correcting postprocessor," *IEEE Trans. Magn.*, vol. 35, no. 5, pp. 4378-4386, Sep. 1999.

[23] T.R. Oenning and J. Moon, "A low-density generator matrix interpretation of parallel concatenated single bit parity codes," *IEEE Trans. Magn.*, vol. 37, no. 2, pp. 737-741, Mar. 2001.

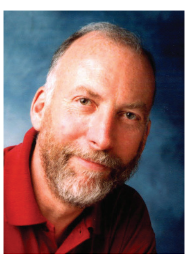
[24] T. Kanaoka and T. Morita, "Structured LDPC codes with reversed MTR/ECC for magnetic recording channels," *IEEE Trans. Magn.*, vol. 42, no. 10, pp. 2561-2563, Oct. 2006.

[25] M. Blaum, R. Galbraith, K. Lakovic, and B. Wilson, "Reverse concatenation with maximum transion run (MTR) codes for high-density perpendicular recording," in *Proc. IEEE Intl. Conf. Global Telecommun.* (GLOBECOM), New Orleans, USA, Nov. 2008, pp. 1-4.

[26] G. Bouwhuis, J. Braat, A. Huijser, J. Pasman, G. van Rosmalen, and K.A.S. Immink, *Principles of Optical Disc Systems.* Adam Hilger Ltd, 1985, Chap. 2.

[27] K. Cai, K.A.S. Immink, Z. Qin, L. Shi, and X. Miao, "Error correction code failure rate analysis for parity-check coded optical recording systems," *Jpn. J. Appl. Phys.*, vol. 47, no. 7, pp. 5870-5874, 2008.

Kui Cai received B.Eng. degree in information and control engineering from Shanghai Jiao Tong University, Shanghai, China, in 1992, M.Eng. degree in electrical engineering from National University of Singapore, Singapore, in 2000, and a joint Ph.D. degree in electrical engineering from Technical University of Eindhoven, The Netherlands, and National University of Singapore, Singapore, in 2007. Since 1999, she has been with Data Storage Institute, Singapore, where she is currently a Senior Research Fellow. Her research interests include coding theory, communication theory, and signal processing for data storage systems.



Kees A. Schouhamer Immink received his PhD degree from the Eindhoven University of Technology. He founded and was named president of Turing Machines Inc. in 1998. He is, since 1994, an adjunct professor at the Institute for Experimental Mathematics, Essen University, Germany, and is affiliated with the Nanyang Technological University of Singapore. Immink designed coding techniques of a wealth of digital audio and video recording products, such as Compact Disc, CD-ROM, CD-Video, Digital Compact Cassette system, DCC, Digital Versatile Disc, DVD, Video Disc Recorder, and Blu-ray Disc. He received a Knighthood in 2000, a personal Emmy award in 2004, the 1996 IEEE Masaru Ibuka Consumer Electronics Award, the 1998 IEEE Edison Medal, 1999 AES Gold and Silver Medals, and the 2004 SMPTE Progress Medal. He was named a fellow of the IEEE, AES, and SMPTE, and was inducted into the Consumer Electronics Hall of Fame, and elected into the Royal Netherlands Academy of Sciences and the US National Academy of Engineering. He served the profession as President of the Audio Engineering Society inc., New York, in 2003.

Yuan Xing Lee received B.S.E.E. degree from Institute of Information Engineering, China, in 1986, and M.E. and Ph.D. degrees in electrical engineering from Xidian University, Xian, China, in 1989 and 1992 respectively. From 1992 to 1999, he was with Data Storage Institute Singapore where he managed Signal Processing and Coding Dept. From 1999 to 2002, he was with IBM Storage Systems Division San Jose. From 2003 to 2006, he was with Hitachi Global Storage Technologies (HGST) San Jose where he managed the Dept of San Jose Signal Processing for Channel Architecture and Dept of San Jose Channel Integration. Since 2007 he has become a senior director of read channel IP development and design in LSI Corporation. Dr. Lee has been involved in advanced signal processing and coding for read/write channels for magnetic recording since 1992. He owns more than 30 US patents and more than 30 journal publications. Dr. Lee is also an active member of IEEE activities. He was a session chair and organizer of many IEEE organized conferences and a secretary of Signal Processing for Storage Committee. He is currently a chair of Data Storage Technical Committee, IEEE Communication Society.

Zhiliang Qin received the B. Eng. and M. Eng. degrees in electrical engineering from Beijing Institute of Technology and Graduate School of China Academy of Engineering Physics, Beijing, P. R. China, in 1995 and 1998, respectively, and the Ph. D. degree in electrical engineering from Nanyang Technological University, Singapore, in 2003. Since Feb. 2002, he has been with Data Storage Institute, Singapore, where he is currently a Senior Research Fellow. His research interests are iterative decoding and digital signal processing over magnetic and optical recording channels, and multiuser detection for code-division multiple-access (CDMA) and multiple-input multiple-output (MIMO) systems.

Tow Chong Chong obtained his B.Eng degree from the Tokyo Institute of Technology, his M.Eng degree from the National University of Singapore, and his Sc.D degree from the Massachusetts Institute of Technology, all in Electrical Engineering. He was awarded the President's Scholarship and the Japanese Government Scholarship for his undergraduate study, and the NUS Overseas graduate scholarship for his graduate study. He also attended the Advanced Management Program at Harvard Business School. He is currently the Executive Director of Science & Engineering Research Council of A*STAR and Executive Director of the Data Storage Institute. Prof Chong's research interest is in the field of magnetic and optical data storage. His other research interests include high-speed electronic and optical devices. Prof Chong is also a Professor with the Department of Electrical and Computer Engineering, NUS. Dr Chong has authored and co-authored more than 200 publications in international refereed journals, presented 23 invited talks and holds 20 patents. He serves as co-chairman of APMRC2002 and as member of the Technical Program Committee for ODS (USA), ISOM (Japan), MORIS (Japan), CLEO Pacific (USA) and OECC (Japan). Prof Chong was awarded the Public Administration Medal (Silver) in the National Day Awards in 2004.